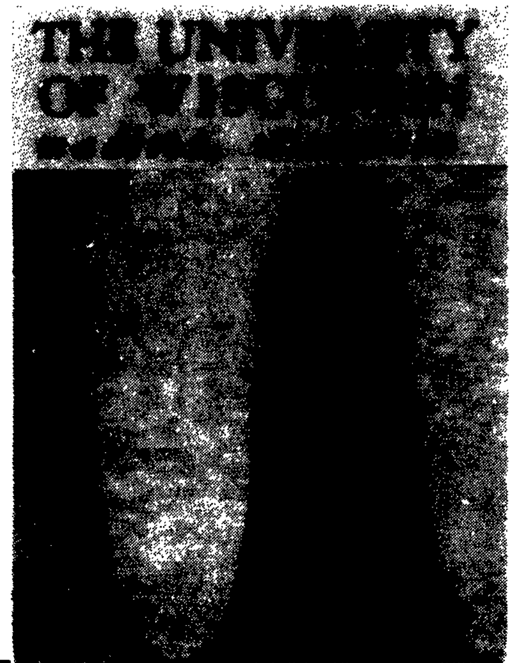
UNCLASSIFIED

63-2-3



MATHEMATICS RESEARCH CENTER

MATHEMATICS RESEARCH CENTER, UNITED STATES ARMY

THE UNIVERSITY OF WISCONSIN

Contract No. DA-11-022-ORD-2059

CONVERGENCE OF RECURRING SERIES WITH GENERAL-ORDER RECURRENCE RELATIONS

F. M. Arscott

MRC Technical Summary Report # 349
December 1962

Madison, Wisconsin

ABSTRACT

The problem considered is that of determining the radius of convergence of the series $\sum_{0}^{\infty} u(r)x^r$, when the coefficients $u(r)$ are connected by a recurrence relation containing $(p + 1)$ terms. Apart from the trivial case $p = 1$, only the case $p = 2$ has hitherto been treated, using the method of continued fractions, which cannot be generalized to $p > 2$.

In this paper, a matrix technique is given, which is applicable to any value of $\underline{p}$, whereby infinite products of $(p \times p)$ matrices are used to formulate characteristic equations and expressions for the $u(r)$. Particular attention is given to the situation in which the problem involves $(p - 1)$ variable parameters, whose values have to be determined in order to secure the greatest possible radius of convergence.

CONVERGENCE OF RECURRING SERIES WITH GENERAL-ORDER RECURRENCE RELATIONS

F. M. Arscott

1. Introduction

In this paper we investigate convergence properties of a series

$$\sum_{0}^{\infty} u(r)x^r \ , \tag{1.1}$$

when there is a recurrence relation between $(p + 1)$ successive coefficients $u(r)$ of the series, of the form

$$u(r+1) = \alpha_1(r)u(r) + \alpha_2(r)u(r-1) + \ldots + \alpha_p(r)u(r-p+1) \ , \quad (r \geq p-1) \tag{1.2}$$

in which the $\alpha_i(r)$ have known asymptotic behavior as $r \to \infty$. The main feature of interest is the phenomenon which I shall call "augmented convergence" ; in general, the series (1.1) will have a certain radius of convergence ρ_1 ; if a certain relation holds between the coefficients $\alpha_i(r)$ and the "starting values" $u(0)$, $u(1), \ldots, u(p-1)$, then the radius of convergence may be increased to a higher value ρ_2 ; if a further relation is also satisfied then the radius of convergence may be increased to a still higher value ρ_3, and so on. Ultimately, if $(p - 1)$ relations are satisfied, the radius of convergence reaches its greatest possible value ρ_p .

Sponsored by the Mathematics Research Center, United States Army, Madison, Wisconsin under Contract No. DA-11-022-ORD-2059.

The investigation of this problem was stimulated by the occurrence of series such as (1.1) in the solution of ordinary differential equations by the method of Frobenius. In the classical equations of hypergeometric type (Legendre's and Bessel's equations, for instance), $p = 1$ so the problem is trivial; D'Alembert's ratio test gives the radius of convergence unambiguous as $|\lim \alpha_1(r)|^{-1}$, and the phenomenon of augmented convergence does not occur. But in equations of higher type than hypergeometric (Mathieu's and Lamé's equations, the spheroidal and ellipsoidal wave equations, etc.) the recurrence relations always involve at least three terms, often four or five. Then, the securing of augmented convergence is equivalent to choosing the parameters in the differential equation so that a solution may converge in a region containing singularities of the equation.

Two forms in which the problem commonly arises are

(i) the $\alpha_i(r)$ of (1.2) are all fully determined functions of $\underline{r}$, while the "starting values" $u(0), \ldots, u(p-1)$ are (initially at least) arbitrary;

(ii) the starting values are given in terms of a single value $u(0)$ by $(p-1)$ supplementary relations of the form

$$\begin{aligned} u(1) &= \alpha_1(0)u(0) \\ u(2) &= \alpha_1(1)u(1) + \alpha_2(1)u(0) \;, \\ &\ldots\ldots\ldots\ldots\ldots\ldots \\ u(p-1) &= \alpha_1(p-2)u(p-2) + \ldots + \alpha_{p-1}(p-2)u(0) \;, \end{aligned} \tag{1.2a}$$

and the functions $\alpha_i(r)$ contain $(p-1)$ parameters. Each of these special cases, and the intermediate situation in which, say, the functions $\alpha_i(r)$

contain $(p - k)$ parameters and there are $(k - 1)$ relations of the form (1.2a), are included in what follows.

Examples of three-term recurrence relations (the case $p = 2$) occurred in the latter part of the 19th century. These were treated by Kelvin and Heine (apparently independently) using a continued-fraction technique which has been extensively developed by Perron (9, Vol. II § 20) Meixner and Schäfke (7, § 1.8) Blanch (5), Bouwkamp (6) and others to the point where it provides a completely satisfactory treatment of the case $p = 2$, both from the theoretical and the computational points of view.

The only drawback to the continued-fraction technique is the apparent impossibility of direct generalization to values of $\underline{p}$ greater than 2. This is hardly surprising, since the essential feature of a continued-fraction is that the successive convegents have the form A_r/B_r, where A_r and B_r each satisfy recurrence relations with precisely three terms. Moreover, a careful study of the technique suggest's that the use of continued fractions provides a convenient formulation of the problem but little positive help towards a solution, since in order to handle such fractions one is thrown back ultimately on to a (possibly different) three-term recurrence relation.

In this paper, therefore, the continued-fraction method has been abandoned; instead we follow the lead given thirty years ago by Milne-Thomson (8, § 5.3) and by George Birkhoff and his co-workers (1-4), and use a matrix formulation of the problem. This provides an alternative method of treating the three-term recurrence relation and one which may well, in view of the ease with which modern computers handle matrix multiplication, prove more suitable for computational purposes. It also generalizes, very easily and naturally, to the case of recurrence relations of any finite order.

The problem as a whole has three inter-related aspects:-

(i) determination of the possible radii of convergence ρ_i of the series,

(ii) formulation of the conditions to be satisfied in order to secure augmented convergence (the <u>characteristic equations</u>),

(iii) determination of the associated values of the coefficients $u(r)$. Sometimes only (i), or (i) and (ii) are required.

In the present paper, we give an outline of the general method in § 2, then in § 3 re-formulate the problem in matrix notation and obtain several different forms of the conditions for augmented convergence. We are then able to justify, in § 4, a working rule by which the possible radii of convergence may be written down almost immediately. Finally, in §§ 5, 6 the evaluation of the coefficients $u(r)$ is considered, a process which yields further forms of the conditions for augmented convergence.

Besides my obvious indebtedness to the authors cited, my warmest thanks are also due to Professor Tomlinson Fort for drawing my attention to the relevance of Birkhoff's work to the problem.

We shall have occasion to use the following results relating to difference equations of type (1.2), due to Birkhoff (2-3), Birkhoff and

Trjitzinsky (4) and Adams (1). Assume that each coefficient $\alpha_i(r)$ can be represented asymptotically, for large r, by a series containing at most a finite number of positive powers of r, i.e. of the form

$$\alpha_i(r) \sim \sum_{j=-k}^{\infty} \alpha_j^{(i)} r^{-j} , \tag{1.3}$$

with real $\alpha_j^{(i)}$. Then for sufficiently large r there exist p linearly independent solutions of (1.2) say, $w_1(r), w_2(r), \ldots, w_p(r)$.

The actual determination of the $w_i(r)$ is in the form of series which are asymptotically valid for large r. In general, each solution $w_i(r)$ is of the form

$$w_i(r) = r^{\mu_i r} \; t_i^{\,r} \; r^{\sigma_i} \; F_i(r) , \tag{1.4a}$$

where $F_i(r)$ is a series asymptotic in r,

$$F_i(r) \sim 1 + \frac{f_1^{(i)}}{r} + \frac{f_2^{(i)}}{r^2} + \ldots . \tag{1.4b}$$

The constants μ, t, σ, f_1, f_2, ... may be found successively by substitution of the formal series (1.4) into the equation (1.2). The determination of μ is normally simple; it is real, rational and in many cases is either zero or an integer which may be found by inspection. Then the possible values of t appear as roots of an algebraic equation of degree not more than p. When μ and t have been evaluated, σ can be found with a little more difficulty, and one can then proceed to the determination of f_1, f_2, ... ; for discussing the convergence properties of $\Sigma\, u_r x^r$, however, we do not need to know these and indeed it is often sufficient to have the values of μ and t.

This, as has been said, is the solution "in general", and here "in general" means provided the value of t is not a multiple root of the equation in which it appears. If this is not the case, however, and we have an N-fold root, then N solutions are obtained which may take different forms according to circumstances - indeed, there are so many possibilities that even a summary would be inordinately long. The case of double roots should be mentioned, however; if we have a value of μ_1 for which a double root t_1 occurs then there are normally two corresponding solutions of the form

$$r^{\mu_1 r} t_1^{\,r} \exp(\pm \gamma r^{\frac{1}{2}})\, r^{\sigma_1} [1 + \frac{\varphi_1}{r^{\frac{1}{2}}} + \frac{\varphi_2}{r} + \ldots.] \qquad (1.5)$$

γ being a further real constant (1). Even this rule, however, has exceptions and logarithmic terms may occur (3).

2. Outline of the method

In order to keep the exposition within reasonable bounds, we shall make here the simplefying assumption that all the solutions $w_i(r)$ of (1.2) are of the form (1.4) or the form (1.5). As has been mentioned above, the cases thus excluded are highly exceptional, and it seems that extension of the method to cover these would not present any intrinsic difficulty.

The solutions $w_i(r)$ can clearly be ordered in such a way that, as $r \to \infty$, we have

$$|w_1(r)| \geq |w_2(r)| \geq \ldots\ldots \geq |w_p(r)| \tag{2.1}$$

and in what follows we assume this has been done.

Now the equation (1.2) is linear and homogeneous, so any solution $u(r)$ must be expressible as a linear combination of the $w_i(r)$, for sufficiently great r - that is to say, there must be constants a_i ($i = 1$ to p) such that

$$u(r) = \sum_{i=1}^{p} a_i w_1(r) \tag{2.2}$$

for r sufficiently great; the constants a_i will depend on, and be uniquely determined by, the values taken by the functions $\alpha_1(r), \ldots, \alpha_p(r)$ and the starting-values $u(0), \ldots, u(p-1)$.

Suppose, first, that the functions $w_i(r)$ are such that $w_j(r)/w_i(r) \to 0$ as $r \to \infty$ for $i < j$; brief consideration of the forms (1.4), (1.5) shows that this will be the case provided we exclude the case $\mu_i = \mu_j$,

$|t_i| = |t_j|$, $t_i \neq t_j$, that is, we do not have two solutions with the same value of μ and values of t which differ in amplitude but not in modulus. Then by writing (2.2) in the form

$$u(r) = w_1(r)[a_1 + \sum_2^p a_i w_i(r)/w_1(r)] \tag{2.3}$$

we obtain immediately the result that as $r \to \infty$,

$$u(r) \sim a_1 w_1(r) , \tag{2.4}$$

provided $a_1 \neq 0$, so that the series $\Sigma u(r)x^r$ and $\Sigma w_1(r)x^r$ are equiconvergent. Application of the Cauchy convergence test then shows at once that the radius of convergence of $\Sigma\, u(r)x^r$ is ρ_1, where

$$\rho_1 = \lim_{r\to\infty} |w_1(r)|^{-1/r} \quad ; \tag{2.5}$$

from (1.4), (1.5) this clearly has the value 0 if $\mu_1 > 0$, ∞ if $\mu_1 < 0$ and $|t_1|^{-1}$ if $\mu_1 = 0$.

If, however, $a_1 = 0$, then (2.4) no longer holds; instead of (2.3) we may write

$$u(r) = w_2(r)[a_2 + \sum_3^p w_i(r)/w_2(r)] , \tag{2.6}$$

giving

$$u(r) \sim a_2\, w_2(r) , \tag{2.7}$$

provided $a_2 \neq 0$, so that the radius of convergence of $\Sigma\, u(r)x^r$ becomes $\rho_2 \equiv \lim |w_2(r)|^{-1/r}$, i.e. 0 if $\mu_2 > 0$, ∞ if $\mu_2 < 0$, $|t_2|^{-1}$ if

$\mu_2 = 0$. This is certainly not less than ρ_1, and will in fact be greater than ρ_1 if $\mu_2 < \mu_1$ or $\mu_2 = \mu_1$ and $|t_2| < |t_1|$. Similarly, if $a_1 = a_2 = 0$, $a_3 \neq 0$, the radius of convergence becomes $\rho_3 \equiv \lim |w_3(r)|^{-1/r}$.

The effect of this is that the radius of convergence of $\Sigma\, u(r)x^r$ will in general have a certain value; if we can choose the starting-values $u(0), \ldots, u_{p-1}$ or parameters in the functions $\alpha_i(r)$ or both, in such a way as to make the constant a_1 vanish, then the radius of convergence may be increased. If we can also make a_2 vanish then the radius of convergence may be increased still further, and so on.

Now consider the case excluded above, namely when two or more of the $w_i(r)$ have the same μ and values of t which are equal in modulus but differ in amplitude. For definiteness, suppose we have $\mu_1 = \mu_2$, $t_1 = Te^{i\theta_1}$, $t_2 = Te^{i\theta_2}$, and $w_i(r)/w_2(r) \to 0$ for $i \geq 3$. Then as $r \to \infty$,

$$u(r) \sim r^{\mu_1 r} T^r \,[a_1 e^{i\theta_1 r} r^{\sigma_1} + a_2 e^{i\theta_2 r} r^{\sigma_2}] \tag{2.8}$$

then somewhat tedious consideration of inequalities show that provided a_1 and a_2 are not both zero,

$$[a_1 e^{i\theta_1 r} r^{\sigma_1} + a_2 e^{i\theta_2 r} r^{\sigma_2}]^{1/r} \to 1\ , \tag{2.9}$$

giving

$$\lim |u(r)|^{-1/r} = \lim_{r\to\infty} r^{\mu_1} T\ , \tag{2.10}$$

so that the radius of convergence of $\Sigma\, u(r)x^r$ is now $\rho_1 \equiv \lim\limits_{r\to\infty} r^{-\mu_1} |t_1|^{-1}$, and this will be the case provided a_1, a_2 do not both vanish. If, however, $a_1 = a_2 = 0$, then the radius of convergence becomes $\rho_3 = \lim |w_3(r)|^{-1/r}$.

The same argument is applied without difficulty to similar cases. If, for instance, we have $w_i(r)/w_1(r) \to 0$ for $i = 2, 3, \ldots, p$, $w_i(r)/w_2(r) \to 0$ for $i = 4, 5, \ldots, p$, but $\mu_2 = \mu_3$, $|t_2| = |t_3|$, $t_2 \neq t_3$ then the radius of convergence is ρ_1 unless $a_1 = 0$, ρ_2 if $a_1 = 0$ and either $a_2 \neq 0$ or $a_3 \neq 0$, and ρ_4 if $a_1 = a_2 = a_3 = 0$, etc.

3. Formal solution of the problem

As remarked above, calculation of the solutions $w_i(r)$ of (1.2), in full, is a task of prohibitive magnitude, and the technique to be employed makes use of a device whereby we only need to determine the dominant parts of the $w_i(r)$, a comparatively simple matter. We write $v_i(r)$ for the part of $w_i(r)$ which is left after omitting the terms which are $o(1)$, and $y_i(r)$ for the part of $v_i(r)$ left after omitting the r^{σ_i} term. That is to say, if $w_i(r)$ is given by (1.4) then

$$v_i(r) = r^{\mu_i r} t_i^r r^{\sigma_i}, \quad y_i(r) = r^{\mu_i r} t_i^r, \tag{3.1a}$$

and if $w_i(r)$ is given by (1.5),

$$v_i(r) = r^{\mu_i r} t_i^r \exp(\gamma r^{\frac{1}{2}}) r^{\sigma_i}, \qquad y_i(r) = r^{\mu_i r} t_i^r \exp(\gamma r^{\frac{1}{2}}), \tag{3.1b}$$

so that as $r \to \infty$,

$$w_i(r) \sim v_i(r) \sim y_i(r)\, r^{\sigma_i} . \tag{3.1c}$$

The problem becomes amenable to treatment only when we turn (1.2) into a matrix equation of the first order. We introduce column vectors $U(r)$, A, and matrices $K(r)$, $W(r)$, $V(r)$, $Y(r)$ as follows:-

$$U(r) = \{u(r+1),\ u(r),\ u(r-1), \ldots, u(r-p+2)\}, \tag{3.2}$$

$$A = \{a_1, a_2, a_3, \ldots, a_p\}, \tag{3.3}$$

$$K(r) = \begin{bmatrix} \alpha_1(r) & \alpha_2(r) & \alpha_3(r) & \ldots\ldots & \alpha_p(r) \\ 1 & 0 & 0 & \ldots\ldots & 0 \\ 0 & 1 & 0 & \ldots\ldots & 0 \\ 0 & 0 & 1 & \ldots\ldots & 0 \\ \ldots & \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & 0 & \ldots\ 1 & 0 \end{bmatrix} \tag{3.4}$$

(Observe that det $K(r) = (-)^{p+1}\alpha_p(r)$ in (3.4))

$$W(r) = \begin{bmatrix} w_1(r+1) & w_2(r+1) & \cdots\cdots & w_p(r+1) \\ w_1(r) & w_2(r) & \cdots\cdots & w_p(r) \\ w_1(r-1) & w_2(r-1) & \cdots\cdots & w_p(r-1) \\ \cdots & \cdots & \cdots & \cdots \\ w_1(r-p+2) & w_2(r-p+2) & \cdots\cdots & w_p(r-p+2) \end{bmatrix}, \quad (3.5)$$

$V(r)$, $Y(r)$ being the same in form as $W(r)$ with w_i replaced by v_i, y_i respectively.

Now each of the sets of functions $w_i(r)$, $v_i(r)$, $y_i(r)$ are linearly independent, so that there can be at most a finite number of values of r for which any of the matrices $W(r)$, $V(r)$ or $Y(r)$ are singular. There is no real loss of generality if we suppose that $W(r)$, $V(r)$, $Y(r)$ are all non-singular for all $r \geq p - 1$, for otherwise we simply ignore a finite number of terms at the beginning of the series, convergence not being affected. Similarly, there is no loss of generality in assuming that (2.2) holds for all $r \geq p - 1$.

This means that not only do we have (as we know already) the constant vector A such that

$$U(r) = W(r)\,A\ , \qquad (3.6)$$

but also uniquely determined vectors $B(r)$, $C(r)$ given by

$$U(r) = V(r)B(r) = Y(r)C(r)\ . \qquad (3.7)$$

Now we have, clearly, as $r \to \infty$,

$$W(r) \sim V(r) \quad ,$$

so that

$$B(r) = [V(r)]^{-1}U(r) = [V(r)]^{-1}W(r)A \sim A \quad . \tag{3.8}$$

Moreover, if we introduce the diagonal matrix $D(r)$ given by

$$D(r) = \langle r^{\sigma_1}, r^{\sigma_2}, \ldots, r^{\sigma_p} \rangle \quad , \tag{3.9}$$

then from the fact that as $r \to \infty$, for fixed $\underline{k}$, $(r+k)^{\sigma_i} \sim r^{\sigma_i}$ we have immediately

$$V(r) \sim D(r)Y(r) \quad ,$$

and hence

$$\begin{aligned} C(r) &= [Y(r)]^{-1}U(r) \sim D(r)[V(r)]^{-1}U(r) \quad , \\ &= D(r)[V(r)]^{-1}W(r)\, A \quad , \\ &\sim D(r)\, A \quad . \end{aligned} \tag{3.11}$$

Now by means of (3.2), (3.4) we reformulate equation (1.2) as

$$U(r) = K(r)U(r-1) \quad , \qquad (r \geq p - 1) \quad . \tag{3.12}$$

the "starting values" $u(0), \ldots, u(p-1)$ constituting the "starting vector" $U(p-2)$. Hence we have

$$U(r+1) = K(r+1)K(r)U(r-1) \quad ,$$

etc., and if we write, for brevity,

$$L(n,s) = K(n+s)K(n+s-1)\ldots K(n+1) \quad , \quad (s \geq 1), \quad L(n,0) = 1 \quad , \tag{3.13}$$

we have

$$U(n+s) = L(p-2, r-p+2)U(p-2) \tag{3.15}$$

which, by expressing $U(r)$ in terms of the known elements of the matrices $K(r)$ and the elements of the starting vector $U(p-2)$, provides a formal solution of the equation (3.12). It tells us nothing, however, about the convergence of the corresponding series $\Sigma\, u(r)x^r$. So we introduce the vector $B(r)$ from (3.7) and have

$$B(r) = [V(r)]^{-1}L(p-2, r-p+2)U(p-2) \quad , \tag{3.16}$$

and (3.8) then gives

$$A = \lim_{r\to\infty} [V(r)]^{-1}L(p-2, r-p+2)U(p-2) \quad . \tag{3.17}$$

So the condition $a_i = 0$ is equivalent to the condition that the $\underline{i}^{th}$ element in $[V(r)]^{-1}L(p-2, r-p+2)U(p-2)$ should tend to zero as $r \to \infty$. To express this in matrix form, let J_n denote a row vector of $\underline{p}$ elements, in which the $\underline{n}^{th}$ element is unity and the rest all zero, e.g.

$$J_1 = [1, 0, 0, \ldots, 0] \quad , \qquad J_2 = [0, 1, 0, \ldots, 0], \text{ etc.} \tag{3.18}$$

then we have

$$a_i = J_i A \quad , \tag{3.19}$$

so the condition $a_i = 0$ is equivalent to

$$\lim_{r\to\infty} J_i[V(r)]^{-1}L(p-2, r-p+2)U(p-2) = 0 \tag{3.20}$$

In this formula, it should be noted, we are not using the full solutions $w_i(r)$ but only their dominant terms $v_i(r)$ which are easily found.

The radius of convergence of $\Sigma\, u(r)x^r$ will thus be ρ_1 in general, but will be ρ_2 if the equation

$$\lim_{r \to \infty} J_1[V(r)]^{-1} L(p-2,\ r-p+2) U(p-2) = 0\ , \qquad (3.21a)$$

is satisfied, and ρ_3 if the condition

$$\lim_{r \to \infty} J_2[V(r)]^{-1} L(p-2, r-p+2) U(p-2) = 0\ , \qquad (3.21b)$$

is also satisfied, and so on.

The simplest non-trivial example of this is that in which $p = 2$, $\rho_1 < \rho_2$, and the coefficients $\alpha_1(r)$, $\alpha_2(r)$ contain a single parameter λ, say, the ratio u_1/u_0 being given either explicitly or in terms of λ. (This is the situation which arises in the study of Mathieu functions, spheroidal wave functions, and Lamé functions, hitherto studied by the continued-fraction technique). The problem is then to determine λ so that the radius of convergence may be augmented from ρ_1 to ρ_2. In such circumstances (3.21a) provides an equation for calculating the possible values of λ and, in the particular cases mentioned, is generally known as the <u>characteristic equation</u>. To keep in line with this practice, we shall call (3.20) the <u>i^{th} characteristic equation</u>; for brevity, denote it by $E_i = 0$. Then, reverting to general values of <u>p</u>, suppose there are $(p-1)$ parameters $\lambda^{(1)}, \lambda^{(2)}, \ldots, \lambda^{(p-1)}$ involved in the problem. The radius of convergence will be ρ_1 in general; in order that it should be augmented to ρ_2 the one relation $E_1 = 0$ must hold among the parameters $\lambda^{(i)}$; if the further relation $E_2 = 0$ also holds the radius of convergence is augmented to ρ_3, etc.; ultimately the radius of convergence can be increased to ρ_p if $p-1$ relations hold between the $p - 1$ parameters.

A further form of the characteristic equations can be obtained, in which we do not even have to form the matrix $V(r)$ but need only the still simpler matrix $Y(r)$. By similar reasoning to that which led to (3.16) we have

$$C(r) = [Y(r)]^{-1} L(p-2, r-p+2) U(p-2) \quad , \qquad (3.22)$$

from which the coefficients $c_i(r)$ can be obtained immediately (in matrix notation, we have $c_i(r) = J_i C(r)$). Now we know that as $r \to \infty$, $c_i(r) \sim r^{\sigma_i} a_i$, so the condition $c_i(r) \to 0$ does not necessarily imply $a_i = 0$. However,

$$\{c_i(r)\}^{1/r} \sim r^{\sigma_i/r} a_i^{1/r}$$

$$\sim 1 \text{ if } a_i \neq 0 \quad ,$$

$$\sim 0 \text{ if } a_i = 0 \quad ,$$

so that

$$\{c_i(r)\}^{1/r} \to 0 \Leftrightarrow a_i = 0 \quad , \qquad (3.23)$$

and the $\underline{i}^{\text{th}}$ characteristic equation can be written

$$\lim_{r\to\infty} \{J_i [Y(r)]^{-1} L(p-2, r-p+2) U(p-2)\}^{1/r} = 0 \quad . \qquad (3.24)$$

Finally, in the specially favourable case when all the ρ_i are different, yet another form of the characteristic equation is available. For in such a case we have, if $a_i \neq 0$,

$$u(r) \sim a_1 w_1(r) = c_1(r) y_1(r) \ ; \tag{3.25}$$

hence

$$\frac{u(r+1)}{u(r)} \sim \frac{c_1(r+1)\, y_1(r+1)}{c_1(r)\, y_1(r)}$$

but $y_1(r+1)/y_1(r) \sim r^{\mu_1 r} t_1 e^{\mu_1}$, so

$$r^{-\mu_1 r} t_1^{-1} e^{-\mu_1} \frac{u(r+1)}{u(r)} \sim \frac{c_1(r+1)}{c_1(r)} \sim \frac{a_1(r+1)^{\sigma_1}}{a_1 r^{\sigma_1}} \sim (1+\frac{1}{r})^{\sigma_1} \sim 1 \ . \tag{3.26}$$

If, however, $a_1 = 0$, $a_2 \neq 0$, $u(r) \sim c_2(r) y_2(r)$ and we find by similar working that

$$r^{-\mu_1 r} t_1^{-1} r^{-\mu_1} \frac{u(r+1)}{u(r)} \sim r^{-(\mu_1 - \mu_2) r} (t_2/t_1) e^{\mu_2 - \mu_1} \ ,$$

and the right hand side of this approaches the limit 0 if $\mu_1 > \mu_2$, (t_2/t_1) if $\mu_1 = \mu_2$. Hence the condition $a_1 = 0$, $a_2 \neq 0 \Leftrightarrow$

$$\begin{aligned} \lim_{r \to \infty} r^{-\mu_1 r} t_1^{-1} e^{-\mu_1} \frac{u(r+1)}{u(r)} &= 0 \quad \text{if} \quad \mu_1 > \mu_2 \ , \\ &= (t_2/t_1) \quad \text{if} \quad \mu_1 = \mu_2 \ . \end{aligned} \tag{3.27}$$

Since $u(r+1) = J_1 U(r)$, $u(r) = J_2 U(r)$, and $U(r)$ is given by (3.15), this provides an alternative form of the first characteristic equation.

4. Direct determination of the possible radii of convergence

From the results in § 3 we obtain easily the following working rule for determining the possible radii of convergence.

In the relation (1.2), divide by the highest power of r occurring, then let $r \to \infty$, and let the limiting form of (1.2) be

$$\omega u(r+1) = \alpha_1 u(r) + \alpha_2 u(r-1) + \ldots + \alpha_p u(r-p+1) \quad , \qquad (4.1)$$

ω, α_1, $\alpha_2, \ldots, \alpha_p$ being constants, not all zero.

Then the possible radii of convergence, ρ_i, are the roots of the equation

$$-\omega + \alpha_1 \rho + \alpha_2 \rho^2 + \ldots + \alpha_p \rho^{p-1} = 0 \;; \qquad (4.2)$$

this equation is to be regarded as essentially of degree $(p-1)$, so that infinite roots (if they occur) must be counted.

The justification of this rule depends on the fact that the existence of a solution of (1.2) of the form (1.4) or (1.5) with $\mu > 0$ implies the existence of a zero root of (4.2), the existence of a solution with $\mu < 0$ implies an infinite root of (4.2), while for solutions with $\mu = 0$ the possible values of t are given by the reciprocals of the finite non-zero roots of (4.2), as one easily sees by actual substitution of the trial solutions (1.4) or (1.5).

Hence the possible radii of convergence of the series $\Sigma\, u(r) x^r$ can be written down at once from the recurrence relation (1.2); this generalizes the rule given by Perron in (9) Vol. II, § 46.

5. Determination of the coefficients

One might think that when the starting-values or the parameters have been determined in such a way as to secure the desired radius of convergence, the problem would be solved in the sense that the coefficients $u(r)$ can be found successively by use of (1.2). However, in the simplest non-trivial case, namely $p = 2$, it has long been known that this is not so (10). Unless we achieve the impossible by determining the exact values of any parameters involved, and making no approximations whatever at any stage, the small errors involved will inevitably build up, re-introducing undesired functions $w_1(r)$, $w_2(r)$, etc. into $u(r)$, and so destroying the augmented convergence.

To avoid this, we need to have expressions which give the $U(r)$ (and hence the $u(r)$) directly. The process of obtaining these leads to yet another form for the characteristic equations.

For simplicity, suppose first that $\alpha_p(r) \neq 0$ for all $\underline{r}$, so that the matrix $K(r)$ is non-singular for all $\underline{r}$; the exceptional case will be dealt with in the next section.

Suppose we wish our $u(r)$ to be a predetermined combination of the $w_i(r)$ - that is to say, the a_i are to have pre-assigned values. (For instance, we may wish to have $a_i = 0$ for $i = 1, 2, \ldots, p - 1$ in order to secure maximal convergence for $\Sigma\, u(r)x^r$.) Now we have, from (3.7), (3.14)

$$L(n, s)U(n) = V(n + s)B(n + s) \quad , \tag{5.1}$$

hence

$$B(n+s) = [V(n+s)]^{-1} L(n,s) U(n) \quad . \tag{5.2}$$

Now let $s \to \infty$; then by (3.8) the left hand side tends to the definite limit $\underline{A}$, so that the limit of the right-hand side certainly exists. At present we are excluding the case $\alpha_p(r) = 0$, so all the $K(r)$ are non-singular, $L(n, s)$ is also non-singular and hence

$$\lim_{s \to \infty} [L(n,s)]^{-1} V(n+s)$$

exists. But from (3.4) we have

$$U(n) = [L(n,s)]^{-1} V(n+s) B(n+s) \tag{5.3}$$

so that on letting $s \to \infty$ we have (since $B(n + s) \to A$),

$$U(n) = \lim_{s \to \infty} [L(n,s)]^{-1} V(n+s) A \quad . \tag{5.4}$$

In passing, it may be noted that we can easily verify directly the fact that $U(n)$, as given by (5.4), does satisfy the difference equation in its vector form (3.12). Observing that

$$L(n,s) = L(n+1, s-1) K(n+1) \quad ,$$

which is an immediate consequence of the definition of $L(n, s)$ in (3.13), we obtain

$$\begin{aligned} U(n) &= [K(n+1)]^{-1} \lim_{s \to \infty} [L(n+1, s-1)]^{-1} V(n+s) A \\ &= [K(n+1)]^{-1} U(n+1) \quad , \end{aligned}$$

so that $K(n + 1) U(n) = U(n + 1)$; which is the same as (3.12).

Going back to relation (5.4), we bring in another expression for $U(n)$, namely that obtained in (3.15), and combining the two we have

$$L(p-2, n-p+2)U(p-2) = \lim_{s\to\infty} [L(n,s)]^{-1}V(n+s)A \quad . \qquad (5.5)$$

In this, $\underline{n}$ may be any integer $\geq p-1$; for $n = p-1$ we have a simpler form

$$U(p-2) = \lim_{s\to\infty} [L(p-2,s)]^{-1}V(p-2+s)A \quad . \qquad (5.6)$$

Let us consider the significance of this equation. Each side is a $\underline{p}$ - element column vector, so that (5.5) or (5.6) contains effectively $\underline{p}$ linear equations. It provides the necessary conditions which must be satisfied in order that the elements a_i of the vector $\underline{A}$ shall have pre-determined values, these conditions involving the $\underline{p}$ elements of the starting vector $U(p-2)$ and the elements of the matrices $K(r)$ which, as remarked already, may contain parameters. If we wish convergence to be augmented from ρ_1 to ρ_2 then we must have $a_1 = 0$, but the remaining a_i are arbitrary, so that (5.5) will provide precisely one condition to be satisfied, which is thus equivalent to the first characteristic equation (3.21a). If we wish convergence to be augmented to ρ_3, we must have $a_1 = a_2 = 0$, so that (5.5) then has in it two conditions, equivalent to the first and second characteristic equations and so on. For maximal convergence, we must have $a_i = 0$ for $i < p$, so that (5.5) provides $(p-1)$ conditions.

6. Case of a singular $K(r)$

We have now to show how the analysis of § 5 is affected when $\alpha_p(r)$ vanishes for some $\underline{r}$, causing the corresponding matrix $K(r)$ to become singular. For simplicity, we shall deal only with the case when there is a single value of $\underline{r}$, say $r = N$, for which $\alpha_p(r) = 0$.

It is still possible, then, to construct a solution having prescribed values of the a_i, by using (5.4) with $n \geq N$. For $n < N$ we must use the equivalent of (3.15), namely

$$U(n) = L(p - 2, n - p + 2)U(p - 2) \tag{6.1}$$

Since this is used only for the calculation of a finite number of $U(n)$, the small errors inevitably introduced will not build up to destroy augmented convergence. The equation (5.5) still holds good provided we have $n \geq N$, and can be used in the same way to obtain the various characteristic equations.

One other possibility, of considerable interest, arises in this case but not in that of § 5. We have, for the column vector $U(N-1)$,

$$U(N - 1) \equiv \{u(N), \quad u(N - 1), \ldots, u(N - p + 1)\}$$

with

$$U(N - 1) = L(p - 2, N - p + 1)U(p - 2) \tag{6.2}$$

Suppose we choose the $(p - 1)$ parameters in the problem so as to make $u(N) = u(N - 1) = \ldots = u(N - p + 2) = 0$; then since $\alpha_p(N) = 0$ we have $u(N + 1) = 0$ also, and hence $U(N) = 0$. But we have

$U(N+1) = K(N+1)U(N)$, etc., and consequently $U(n) = 0$ for all $n \geq N$. The series (1.1) then reduces to a polynomial of degree $(N - p + 1)$ in $\underline{x}$. Since a polynomial is, of course, convergent everywhere this may, in certain circumstances, provide a means of augmenting the convergence from ρ_p (if this is finite) to infinity.

As an illustration of this, consider the spheroidal wave equation

$$(1 - x^2)\frac{d^2w}{dx^2} - 2x\frac{dw}{dx} + \{\lambda - \mu^2(1 - x^2)^{-1} + \gamma^2(1 - x^2)\} w = 0 \;; \tag{6.3}$$

regarding γ^2 as fixed and λ, μ as disposable parameters. Assuming a solution of the form

$$\sum_0^\infty u(r)x^r \tag{6.4}$$

the general recurrence relation is found to contain four terms (i.e. $p = 3$) and to be

$$\begin{aligned}(2r+1)(2r+2)u(r+1) = (8r^2 + \mu^2 - \lambda - \gamma^2)u(r) + \\ + \{\lambda + 2\gamma^2 - (2r-1)(2r-2)\}\, u(r-1) + \gamma^2 u(r-2) \;.\end{aligned} \tag{6.5}$$

Dividing through by $4r^2$ and letting $r \to \infty$ we obtain, in the notation of § 4,

$$\omega = 1, \quad \alpha_1 = 2, \quad \alpha_2 = -1, \quad \alpha_3 = 0 \;, \tag{6.6}$$

giving $\rho_1 = \rho_2 = 1$, $\rho_3 = \infty$. Hence by suitable choice of λ and μ we can obtain an infinite series solution which is an integral function of $\underline{x}$. Here $\alpha_p(r) = \gamma^2 \neq 0$ so the general method of § 5 applies.

Now, however, consider the degenerate case where $\gamma = 0$ (i.e. the

Associated Legendre equation). The situation is now different, since we have $p = 2$, and the possible radii of convergence are 1, 1 so that a solution which is an <u>infinite</u> series of the form (6.4) cannot possibly converge outside $|x| = 1$. The only way to obtain a solution with better convergence properties is to make use of the fact that now, since $p = 2$, we have $\alpha_p(r) = \lambda - (2r-1)(2r-2)$, so that if we choose λ as $(2N-1)(2N-2)$, $\alpha_p(N) = 0$ and we have the situation of this paragraph. By suitable choice of μ we can make $u(N)$ vanish, and the solution then becomes a polynomial of degree $(N-1)$ in x^2.

REFERENCES

1. Adams, On the irregular case of the linear ordinary difference equation, Trans. Am. Math. Soc. 30, 1928, 507-541, 855.
2. Birkhoff (Geo. D.), General theory of linear difference equations, ibid. 12, 1911, 243-284.
3. Birkhoff (Geo. D.), Formal theory of irregular difference equations, Acta. Math. 54, 1930, 205-46.
4. Birkhoff (Geo. D.) and Trjitzinsky, Analytic theory of singular difference equations, ibid. 60, 1933, 1-89.
5. Blanch, On the computation of Mathieu functions, J. Math. Phys., 26, 1946, 1-20.
6. Bouwkamp, On spheroidal wave functions of order zero, ibid, 26, 1947, 79-92.
7. Meixner and Schäfke, Mathieusche Funktionen und Sphäroidfunktionen, Springer, 1954.
8. Milne-Thomson, Calculus of finite differences, Macmillan, 1933.
9. Perron, Die Lehre von den Kettenbruchen (3rd edn.) Teubner, 1957.
10. Goldstein, Mathieu functions, Trans. Phil. Soc. Cambridge, 23, 1927, 303-36.